NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

JOINT APPLIED PROJECT

Assessment on
How Much DoD Information Technology Testing Is Enough

By: Kelvin L. Graves

December 2010

Advisor: Brad R. Naegle
Second Reader: Karl Pfeiffer

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** December 2010	**3. REPORT TYPE AND DATES COVERED** Joint Applied Project

4. TITLE AND SUBTITLE Assessment on How Much DoD Information Technology Testing is Enough	**5. FUNDING NUMBERS**
6. AUTHOR(S) Kelvin L. Graves	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) Business Transformation Agency	**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number: N/A.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited	**12b. DISTRIBUTION CODE**

13. ABSTRACT (maximum 200 words)

The Department of Defense has mandated that all military programs implement integrated testing within their program whenever possible; however, the concern of information technology being agile steadily grows within the Department. The Chairman of the House Armed Service Committee appointed the Defense Acquisition Reform Panel on March 2009, to review the acquisition processes. The panel study claimed that the nature of defense acquisition has considerably changed over the years however; the acquisition process has not been able to keep up with the changes. Moreover, the current position on how DoD buys, adopts, creates and implements testing of information technology has become an antiquated process. DoD is looking for an agile process to rapidly develop and deploy information technology to the user while the technology is still relevant. The purpose of this research is to determine how much information technology testing is deemed necessary prior to deploying equipment to DoD users.

14. SUBJECT TERMS Information Technology, Testing and Evaluation and Business Transformation Agency, Business Systems	**15. NUMBER OF PAGES** 66
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for Public release; distribution is unlimited

ASSESSMENT ON HOW MUCH DOD INFORMATION TECHNOLOGY TESTING IS ENOUGH

Kelvin L. Graves
Major, United States Army
B.S., University of Mobile, 1998

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN PROGRAM MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
December 2010

Authors: ____________________________________
Kelvin L. Graves

Approved by: ____________________________________
Brad R. Naegle, Lead Advisor

____
Dr. Karl Pfeiffer, Support Advisor

____
William R. Gates, Dean
Graduate School of Business and Public Policy

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The Department of Defense has mandated that all military programs implement integrated testing within their program whenever possible; however, the concern of information technology being agile steadily grows within the Department. The Chairman of the House Armed Service Committee appointed the Defense Acquisition Reform Panel on March 2009, to review the acquisition processes. The panel study claimed that the nature of defense acquisition has considerably changed over the years however; the acquisition process has not been able to keep up with the changes. Moreover, the current position on how DoD buys, adopts, creates and implements testing of information technology has become an antiquated process. DoD is looking for an agile process to rapidly develop and deploy information technology to the user while the technology is still relevant. The purpose of this research is to determine how much information technology testing is deemed necessary prior to deploying equipment to DoD users.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

ACAT	Acquisition Category
AFOTEC	Air Force Operational Test and Evaluation Center
ATEC	Army Test and Evaluation Command
BCL	Business Capability Model
BTA	Business Transformation Agency
CDD	Capability Development Document
CIO	Chief Information Officer
CJCSI	Chairman of the Joint Chiefs of Staff
CPD	Capability Production Document
CT&E	Capability Test and Evaluation
DAA	Designated Accrediting Authority
DAB	Defense Acquisition Board
DAR	Defense Acquisition Reform
DAS	Defense Acquisition System
DAU	Defense Acquisition University
DBSMC	Defense Business System Management Committee
DISA	Defense Information System Agency
DoD	Department of Defense
DoDI	Department of Defense Instructions
DSB	Defense Science Board
ERAM	Enterprise Risk Assessment Methodology
GAO	Government Accountability Office
GV&V	Government Validation and Verification
HASC	House Armed Services Committee
IA	Information Assurance
ICD	Initial Capability Document
IRB	Investment Review Board

IT	Information Technology
JCIDS	Joint Capabilities Integration and Development System
JITC	Joint Interoperability Test Command
MCOTEA	Marine Corps Operational Test and Evaluation Activity
MDA	Milestone Decision Authority
NSS	National Security Systems
OA	Operational Assessment
OPTEVFOR	Operational Test and Evaluation Force
OTAs	Operational Test Agencies
PEO	Program Executive Officer
PM	Program Manager
PPBE	Planning, Programming, Budget, and Execution
SPS	Standard Procurement System
SPS-B	Standard Procurement System-Bugzilla
T&E	Test and Evaluation
WR	Warfighter Requirement Directorate

ACKNOWLEDGMENTS

I would like to express my thanks to my advisors, Mr. Brad Naegle and Dr. Karl Pfeiffer for providing guidance and assistance with this project. Their encouragement and advice throughout this process has been invaluable. There are no words to express my gratitude for your time and support. I would also like to thank the NPS Acquisition Research Program and the NPS Thesis Process office for your support during this process.

My family wife, Thayda, and sons George and William are the pillars that have supported me throughout my 18 months at NPS. They always gave the right push when I was exhausted that motivated me to finish as well as I had started in the beginning of this journey. Their devotion, support and inspiration gave me the strength to finish strong. I am extremely grateful and blessed to have you in my life. I love you.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. PURPOSE

The purpose of this research is to examine how much information technology (IT) testing is necessary prior to deploying information technology equipment to Department of Defense (DoD) users. The intent is to determine if a new acquisition process for information technology is suitable and feasible in reducing the current schedule required to provide new technology to the DoD warfighters. This researcher reviews DoD information technology reforms, DoD House Armed Services Committee (HASC) reports, and the Weapon System Acquisition Reform Act, along with other literature, in search of the best solution for the Department of Defense. In ascertaining which acquisition model might work best, the researcher explores the IT testing process and examines why the DoD is pursuing a new model for acquiring and testing IT. The objectives of this research are threefold: assess whether it is feasible and in the best interests of the DoD to implement a new testing model for information technology; examine what an improved model should depict; and assess the associated risk of using an agile approach to testing IT.

B. BACKGROUND

The Department of Defense has mandated that all military programs implement integrated testing within their programs whenever possible. However, there is growing concern among acquisition professionals that information technology is more effectively developed using agile methodologies. According to Burton, Hammons, Lapham, Schenker, and Williams (2010), agile is defined as an iterative and incremental (evolutionary) approach to software development, which is performed in a highly collaborative manner by self-organizing teams (p. 5). The Chairman of the House Armed Services Committee appointed the Defense Acquisition Reform Panel in March 2009 to review the acquisition processes. The panel study (Andrews et al., 2010) concluded that the nature of defense acquisition has considerably changed over the years; however, the

acquisition process has not been able to keep up with these changes. Moreover, the current acquisition process controlling how the DoD buys, adopts, creates, and implements testing on information technology has become an antiquated process. The DoD is examining an agile process to rapidly develop and deploy information technology to users while the technology is still relevant.

The impact of today's economy and the use of federal funds have become reasons for Congress to question every acquisition program. As stated in a Congressional Research Service report, the structure of the DoD acquisition system has been a concern of Congress for many years (Schwartz, 2009, p. 1). Moreover, program managers are faced not only with meeting the user requirements but also with budgetary restrictions. As the need for technology increases, how does the DoD provide technology to warfighters in a timely manner within the DoD's current acquisition model?

Information technology has revolutionized the way the DoD has conducted wars, deployments, and conflicts over the centuries. However, the acquisition system that the DoD is currently using for information technology is antiquated in providing the latest technology to warfighters. In the Business Executives for National Security (BENS) cooperative report to the Defense Information Systems Agency (DISA), it stated that "Congress, DoD and its industry partners have constructed a complex acquisition system designed first and foremost to promote fairness and prevent abuse" (Business Executives for National Security [BENS], 2008, p. 3). It further asserted that "unlike its commercial counterparts, which emphasize time-to-market and competition, the DoD system (in fact, all of Federal procurement) is process driven and encrusted with a statute and regulatory-driven organizational structure that confuses oversight with management review" (BENS, 2008, p. 3).

The DoD finds itself still using one model for the entire acquisition of systems, which has proven that it slows the rate for new information technology development and deployment. This does not mean that its current model has not provided quality systems from the acquisition process, but the timeliness of this process has impacted the warfighters capabilities during previous conflicts and wars. Deputy Defense Secretary

William J. Lynn III stated in a press conference that the DoD has the best weapons systems the world has ever seen, but in the IT area, “our system has followed that model, and it simply doesn’t work” (Garamone, 2010).

The acquisition process has some key gates through which all systems must pass in order to be fielded to the intended users. The process that these systems must undergo can be shortened based on the need, technical maturity, and type of equipment that users are requiring in a particular environment. In this report, the researcher analyzes the acquisition process briefly and examines the relationship between information technology and DoD testing.

C. SCOPE

This research has four objectives. The first is to assess whether it is feasible and in the best interest of the DoD to implement a new testing model. Multiple acquisition models are available for tailoring the acquisition process to provide a more rapid response in developing the system or product and deploying it to the users. This researcher evaluates the current weapon system testing cycle against other information technology models that are proposed. On the basis of these models, this researcher analyzes those approaches along with how they might improve the information technology development process.

Second, this researcher assesses the impact that the Business Capability Model has provided in developing business systems and deploying them to the warfighter. The Business Transformation Agency (BTA) has created this model in search of a faster approach to getting technology into the user’s hands. The researcher evaluates the possible courses of action for reducing the duration of information technology testing for Acquisition Category III IT systems by examining the Business Capability Lifecycle (BCL) model proposed by the BTA.

Third, the researcher assesses historical testing data with regard to the number of defects discovered and how historical defect data can help determine how much testing is sufficient to reduce risks to acceptable levels. Because there is no way to eliminate all defects within software testing, the question remains of when to stop testing and deploy

the product to the user. This portion of the research is focused on defect removal efficiency and how it impacts future information systems testing. The defects of a system under testing play a major factor in when a system is released to the users. Also, risk cost and schedule dictate the release of a system. Currently, operational assessment provides programs with the ability to rapidly deploy an operationally effective system versus delaying the system deployment until it has completed the entire test program.

Finally, the researcher closely examines the associated risk of using an agile approach to testing information technology. The intent is to determine and assess the implications of testing and deploying IT faster to DoD users. The overall objective of this research is to answer the following questions.

1. Primary Research Question

- On an ACAT III IT program, how much testing is required to ensure that risk is minimized to an acceptable level?

2. Secondary Research Questions

- How does previous software testing such as Government Validation and Verification (GV&V) Testing, Regression Testing, and System Acceptance Testing compare and impact the final decision to deploy a system to users?
- How much additional testing is worth the cost and time for ACAT III programs?

D. RESEARCH METHOD

The methodology used in this research included data collection from the Standard Procurement System (SPS) program office, which included test reports, SPS-Bugzilla reports, and customer service reports. Additional data came from the BTA acquisition process. Data analysis was conducted to identify additional problems in order to understand the implication of how much testing is truly required for business systems

within the Department of Defense. As part of the data collection effort, the researcher also analyzed the different acquisition models that are currently used and the Defense Science Board–Information Technology (DSB–IT) proposed model in expediting the release of new technology faster.

E. CHAPTER OVERVIEWS

The researcher explores the current acquisition process, system engineering process, and SPS data, along with GAO reports and other literature reviews, to determine how much information technology testing is sufficient to exit the testing phase. In Chapter I, the researcher introduced the reader to the purpose and scope of this research. The researcher also provided the foundation for the reader to understand the importance of this research.

Background on the acquisition process begins in Chapter II, along with the testing aspect of the DoD. The intent is to lay the foundation for the remaining chapters that move more into ACAT III IT systems and how the processes could improve to decrease the time and cost of all business systems. In Chapter III, the researcher introduces the BTA along with its acquisition model and governance processes. In this chapter, the researcher also introduces the data from the Standard Procurement System program and introduces how the analysis is conducted in Chapter IV. The intent of Chapter IV is to conduct an analysis of the data introduced in Chapter III. Finally, the researcher concludes Chapter V by answering the primary and secondary research questions, along with giving recommendations for future research topics on testing information technology.

II. IT ACQUISITION AND TESTING BACKGROUND

A. INTRODUCTION

The DoD has been acquiring military information technology equipment for decades with the intent of maintaining information dominance over its enemies. Like any corporation, the more market shares it has in the global world, the more powerful it is within that environment. The DoD acquisition process has provided that competitive edge for its warfighters throughout the advancement of technology. However, as with any process, bottlenecks can occur. To understand how to resolve the problem, one most examine the current process and determine if there are no-value-added processes within the system. The no-value-added process must be eliminated to reestablish a lean process that produces effective results.

1. Current Acquisition Process

The process of information technology acquisition is typically controlled through the Defense Acquisition System (DAS) from initiation (becoming a program of record) to eventual disposal (reaching the end of its lifecycle). The DAS management process also works parallel with other processes including the Joint Capability Integration Development System (JCIDS) and the Planning Programming, Budgeting, and Execution (PPBE) system. Within any process, gates and gatekeepers maintain a certain portion of the management process. These gatekeepers for the DAS are Milestone Decision Authorities (MDAs) that ensure everything is progressing according to the acquisition strategy plan. The DAS governing document is the DoD 5000.01, *The Defense Acquisition System* (Under Secretary of Defense, Acquisition, Technology, & Logistics [USD(AT&L)], 2007), which provides the policies that governs the way the process is implemented. Figure 1 is an excerpt from DoD Instruction (DoDI) 5000.02 (USD[AT&L], 2008) and depicts the framework that drives the acquisition process. In Figure 1, the gates are milestones, and, within the gates, additional process owners such as program managers, test agencies, material developers, and other key organizations are

embedded throughout the acquisition process. These organizations impact the acquisition process timeline. As Hutchinson (2009) wrote, "suffice it to say that in the course of creating the acquisition process, we have built a complex environment of rice bowls (meaning a person's small part of a bigger process) and process ownership" (p. 16).

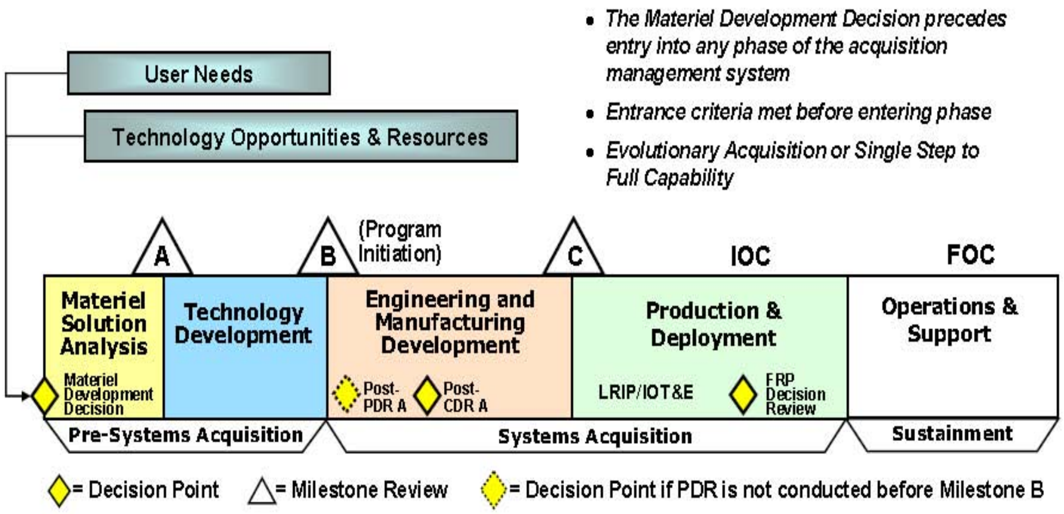

Figure 1. FiguDoD Acquisition Management System (USD[AT&L], 2008, p. 12)

2. Milestones

The management of a complex system, whether it is a weapon or automated information system, can lead to a failure during the testing phases. Within the acquisition system, there are three distinct phases that systems must go through in a serial process. The three phases are pre-system acquisition, system development and demonstration, and production and deployment, as depicted in Figure 1. During these three phases, numerous documents are produced. Coordination of these documents must occur for programs to move forward in producing a product for DoD users. Figure 1 depicts where the decisions points are located throughout the milestones, along with the program initiation and how weapon systems, along with information systems, move throughout the acquisition phases to reach an initial operation capability and full operational capability decisions.

The key documents that are produced during these phases are an Acquisition Strategy, Test, and Evaluation Master Plan (TEMP); Initial Capability Documents (ICDs); Capability Development Documents (CDDs); and a Capability Production Document (CPD). Concept Decision; Milestones A, B, and C; Critical Design; and Full-rate Production Decision Reviews are all key management events that are built within the acquisition process in order to produce a successful system. The milestones, put simply, are decision points that give program managers and decision-makers the ability to review acquisition programs, monitor and administer progress, identify problems and make corrections (Defense Acquisition University [DAU], 2005, p. 50). This current process does not allow for fielding information technology systems to the warfighter in a timely manner.

As stated within the DAU's *Defense Acquisition Guide* (2005), the lifecycle process can be adjusted to fit a particular program, which is normally referred to as "tailoring." (p. 50). As the guidebook states: "The number of phases, key activities, and decision points are tailored by the program manager based on an objective assessment of the program's technical maturity and risks, and the urgency of the mission need" (DAU, 2005, p. 50).

a. Milestone A

The Milestone A decision point occurs during the pre-system acquisition phase. There are two phases within the pre-system acquisition phase that must occur prior to Milestone A: *concept development* and *concept refinement*. During the concept development phase trade-off studies, the analysis of alternatives and the establishment of functional requirements form the Initial Capabilities Document. The concept refinement phase produces program manager's charters, integrated product team's charters, and technology development strategies. The concept refinement phase focuses on innovation, competition, and commercial off-the-shelf equipment.

The technology development phase does not occur until after Milestone A. According to the DAU (2005), the purpose of this phase is to reduce the technology risk in the planned program and to determine the appropriate set of technologies to be

integrated within the system. This phase normally produces prototypes that are placed in a relevant environment to demonstrate their usefulness to the military. The technology development phase leads into the next milestone.

b. Milestone B

After the Milestone B decision, the next phase in the serial process is the Engineering and Manufacturing Development (EMD) phase. During this phase, various system engineering development, integration, and interoperability tests are conducted to develop the product. The critical design review (CDR) is conducted during the EMD phase, and the system must demonstrate its usefulness to the military as one of the exiting criteria out of the EMD phase, which leads to the next milestone.

c. Milestone C

During this phase, operational tests and evaluations are conducted to ensure that the product and system production capability can satisfy the mission requirements. The MDA has the option to either deny or approve a program for low-rate initial production (LRIP). When a program is approved for LRIP, program managers work toward getting a full-rate production decision and producing sufficient systems for initial operational capability (IOC). Typically, the production of LRIP systems is necessary to perform the critical Initial Operational Test and Evaluation (IOT&E) or Operational Evaluation (OPEVAL) using "production representative" systems. Too often, systems have entered into IOT&E or OPEVAL without significant operational-type testing events earlier in the development. The result has been that many systems have failed this critical test after production of the system has begun, and the DoD has taken note. Testing communities are now required to conduct integrated testing when deemed necessary. In order for a system to have a chance to succeed, operational testers must be involved earlier in the developmental process.

d. DoD 5000 Series

The DoD 5000 series includes DoD Directive 5000.01 (USD[AT&L], 2007), DoD Instruction 5000.02 (USD[AT&L], 2008), and other specific governing documents. DoD 5000.01 provides the directive that establishes the Defense Acquisition System. DoDI 5000.02 provides the instruction for implementing the DoD guidance. The Defense Acquisition Guidebook (DoD, 2009) provides additional guidance on best practices that should be tailored to acquisition programs. These documents are the foundation of the DoD Acquisition Management System. Automated Information Systems were included in the DoD 5000 series in the late 1990s. This change brought the weapon system and the automated information system under the same set of rules. Although this made sense to DoD leaders during that time, Congress and acquisition leaders now perceive that a shift needs to occur to reduce the time to fielding for IT systems. The Defense Science Board–Information Technology (DSB–IT) 2009 congressional reports, along with other reports, have pointed this out. This is illustrated by the study conducted by the House Armed Services Committee Panel on Defense Acquisition Reform (DAR) that was established to review the acquisition process:

> As a result, the Department is unable to keep pace with the rate of IT innovation in the commercial market place, cannot fully capitalize on IT-based opportunities, and seldom delivers IT-based capabilities rapidly. By way of example, the private sector is able to deliver capabilities and incrementally improve on those initial deliveries on a 12 to 18 month cycle; defense IT systems typically take 48-60 months to deliver. In an environment where technology is obsolete after 18 months, defense IT systems are typically two to three generations out of date by the time they are delivered. With the exception of IT purchased via vehicles like Enterprise Software Initiative contracts, COTS technologies are insufficiently leveraged, excessively tailored, inefficiently tested and delayed. (DSB–IT, 2009 p. 17)

It is clear from the literature on the DoD 5000 series that other information technology issues are not fully defined within the roles of decision-makers nor in the process dealing with information systems. Steve Hutchison, Test and Evaluation

Executive for Defense Information System Agency, and George Axiotis, from the office of Director, Operational Test and Evaluation Command, noted these key examples:

- Joint Staff J6, and the Designated Accrediting Authority (DAA), respectively, do not sign the T&E master plan (key process owners for interoperability and information assurance), even though they are principal customers of significant T&E activities (Hutchison, 2009b, p. 16).
- The MDA can make a decision to buy IT capability, and the DAA can deny the operation of that capability on the network (Hutchison, 2009b, p. 16).
- "There is no way to establish a baseline on technology that is always changing" (Axiotis, 2009, p. 474).

e. Acquisition Management Roles

To fully understand the acquisition processes, the acquisition management role is reviewed in this section. As mentioned previously, these managers play a crucial role throughout the acquisition process. These managers are held responsible to control the cost, schedule, and performance of a product's lifecycle from the conceptual phase through the disposal phase. Even with this huge responsibility, a program manager is given no power in controlling different organizations that help determine the state of the program. Every program is managed with these key elements: cost, schedule, and performance. The acquisition managers must ensure that programs are meeting the requirements that are set forth in the acquisition strategy prior to transitioning to the next phase (Figure 1).

Program managers, who are responsible for the overall development, production, and deployment of the system, are also responsible for ensuring that user requirements are met. In order to meet these requirements, program managers depend on contractors, the warfighters for the system, the DoD testing communities, and other supporting government agencies in providing support throughout the decision process and milestone reviews. This leads to the next topic, acquisition categories.

f. Acquisition Categories

The system's acquisition category (ACAT) is based on the overall cost of the program or how important the program is viewed by the decision-makers. There are five distinct sections within the ACAT: major defense acquisition programs, major automated defense acquisition programs, major systems, and all other systems (except for the Army, Navy, and Marines Corps, which do not meet the ACAT I- III criteria). These distinct sections are further subdivided into acquisition category numbers from I–IV. To further explain the acquisition categories, Table 1 depicts the ACATs and explains why a program is under that category along with the MDA (DAU, 2005, p. 12). It is also important to note that the dollar value assigned to a program designates its category. Decision-makers can also raise the category on a program with a lesser value to maintain oversight on that program.

Table 1. Acquisition Categories
(DAU, 2005, p. 23)

Major Defense Acquisition Programs	ACAT ID:	• Designated by USD(AT&L) • Defense Acquisition Board Review • Decision by USD(AT&L)	$365M RDT&E or $2.19B Procurement (FY 2000 Constant $)
	ACAT IC:	• Designated by USD(AT&L) • Component-level Review • Decision by Component	
Major Automated Information Systems Acquisition Programs	ACAT IAM:	• Designated by DoD Chief Information Officer • Information Technology Acquisition Board Review • Decision by DoD Chief Information Officer	$378M Life Cycle Cost or $126M Total Prog. Cost or $32M Prog. Cost in any Single Year (FY 2000 Constant $)
	ACAT IAC:	• Designated by DoD Chief Information Officer • Component-level Review • Decision by Component Acquisition Executive	
Major Systems	ACAT II:	• Designated by Component Acquisition Executive • Component-level Review • Decision by Component Acquisition Executive	$140M RDT&E or $660M Procurement (FY 2000 Constant $)
All Other Systems (Except for Army Navy, Marine Corps)	ACAT III:	• Designated IAW Component Policy • Does Not Meet Criteria for ACAT I, IA, II, or III • Review and Decision at Lowest Appropriate Level	No Fiscal Criteria
Army Navy Marine Corps	ACAT IV:	• Designated IAW Component Policy • Does Not Meet Criteria for ACAT I, IA, II, or III • Review and Decision at Lowest Appropriate Level	See AR 70-1 (Army) & SECNAVINST 5000.2C (Navy and Marine Corps)

B. TEST COMMUNITIES OVERVIEW

Five Operational Test Agencies (OTAs) operate within the DoD: Army Test & Evaluation Command (ATEC), Air Force Operational Test Command (AFOTEC), Navy Operational Test and Evaluation Force (OPTEVFOR), Marine Corps Test and Evaluation Activity (MCOTEA), and Joint Interoperability Test Command (JITC). The purpose of OTAs is to test and evaluate systems while employed in missions that are as realistic as possible in order for MDAs to assess whether a program should proceed within the process. OTAs are required by *United States Code* title 10 to conduct Initial Operational Test and Evaluation (IOT&E) on major defense programs.

Title 10 law states that a Major Defense Acquisition Program (MDAP) may not proceed beyond an LRIP until an IOT&E is completed. Testing is an integral component of the acquisition management framework, which determines if a system is performing sufficiently to move forward through the acquisition model. Program Offices (PO) and the OTAs typically have some type of friction that creates problems during the testing process. The test community often tests something that was already tested by another testing agency, which causes redundant data collection efforts and also increases the testing time line. In its cooperative review 2008 report, the Business Executives for National Security (BENS) found that processes were adding to cost and time spent.

> The Developmental Operational Test & Evaluation (DOT&E) and the Operational Test & Evaluation (OT&E) processes, while valuable and necessary, are at times excessive, non-standard across the services, and not particularly well-suited to evaluating software and IT applications, all of which adds to time spent on compliance and adding to cost. (2008, p. 3)

George Axiotis's article about establishing a new acquisition model states that "the DoD cannot wait for optimal solutions before fielding capabilities or rely solely on T&E as its gatekeeper" (2009, p. 478). According to a report submitted to Timothy Harp, the Deputy Assistant Secretary of Defense Command, Control, Communication, Intelligence, Surveillance and Reconnaissance (C3ISR) and IT Acquisition, that

addressed industry perspectives on the future of DoD IT acquisition, there is a need for a new model that requires a combination of architected and agile approaches (Alvidrez et al., 2010, p. vii).

The lingering question that is always addressed by the acquisition community is how much testing is enough. This question depends on whether the developers have the right requirements to design the system and on whether the testers have the right tools and test player to adequately test the system. The majority of the acquisition programs within the DoD are based on some form of software. Testing software for the DoD comes in the form of vendor testing, developmental testing, interoperability testing and operational testing. This testing pattern is done in a serial fashion, and in most cases without conducting integrated testing. As stated in Conley's (2009) article, A test would be performed, data gathered, and then the system would move to the next test center. This process is time consuming, inefficient, and insufficient for network-enabled systems. (p. 111). These test events produce test reports that are given to decision-makers to determine whether a program should proceed to the next phase of the acquisition process.

There are several testing activities (i.e., funding requirements, establishing test units, developing test plans and conditions) that must be met prior to conducting DoD testing. Those conditions come in the form of laws, regulations, and additional testing requirements. DoD testing comes in the form developmental testing and evaluation, operational testing and evaluation, interoperability testing and certification, and information certification and accreditation. Table 2 depicts the roles and condition within these activities.

Table 2. Test and Evaluation in the DoD Acquisition Process (Hutchison, 2009b, p. 18)

Activity	**Test Agent**	**Conditions**	**Customer**	**Reference**
DT&E	PMO/contractor	As determined by PMO; generally benign; lab developer personnel	PMO	DoD 5000
OT&E	Independent OTA	“operational realistic … typical users”	MDA	Title 10 DoD 5000
Joint Interoperability Test Certification	JITC	“applicable capability environment”	J6	DoDD 4630.5 DODI 4630.08 CJCSI 6212.01D
IA Certification and Accreditation (Security T&E) (DIACAP)	OTA, DIA, FSO, NSA	Operational, lab	DAA	DoDI 8510.01 *Note also the DOT&E policy on testing IA during OT&E. DIACAP C&A does not complete the requirement for IA testing

1. Developmental Test and Evaluation

Developmental test and evaluation (DT&E) is designed to help the program manager (PM) fix problems and identify issues early in the program, prior to going to an operational test. During the developmental test, the PM can set the condition for the DT&E environment. DT&E is the primary governmental T&E that verifies that the contractor has met the performance specifications detailed in the contract. The PM typically looks more at the technical side of the system, meaning the functionality of the system, from a system engineering perspective.

2. Operational Test and Evaluation

Operational Test and Evaluation (OT&E) of systems is an independent test conducted by OTAs. OT&E is designed to ensure that the system is operationally effective and operationally suitable considering the warfighter's Doctrine, Organization, Training, Materiel, Leadership and Education, Personnel, and Facilities (DOTMLPF) factors. The OTA has the responsibility of providing relevant feedback to the acquisition decision-makers, giving them information that will support the decision to field a particular system to the warfighter.

3. Interoperability Testing

Interoperability provides the testing required to ensure today's program system can communicate flawlessly with emerging and legacy equipment. JITC is the only OTA with the mandate and authority to certify DoD IT and National Security Systems (NSS) as meeting joint operation requirements (Herrin, Knodle, Mackenzie, & Stephens, 2008, p. 148).

4. Information Assurance Certification and Accreditation

All programs with joint interoperability requirements must go through the certification and accreditation process, along with the supportability and validation process (DAU, 2005, p. 47). This is why it is important for JITC and other agencies to be involved early in the process in order to ensure that systems can get to the warfighters at

a faster pace. When JITC and other certification processes are not embedded in the program, a system can get through the process, but it will not be able to operate within the network. All certification occurs in labs as well as in a realistic environment.

C. CURRENT TESTING TIME LINE

The current testing time line within the DoD, particularly for IT systems, is impacted by the current weapon system model. With the rapid pace of changing technology and evolving threats, the typical testing time line does not allow the DoD to move systems through the systems development process fast enough for the warfighter. The DSB–IT reported in March 2009 that a new model for information technology acquisition is required. Figure 2 depicts the current testing time line, which can be greater than six months (Hutchison, 2010, p. 25).

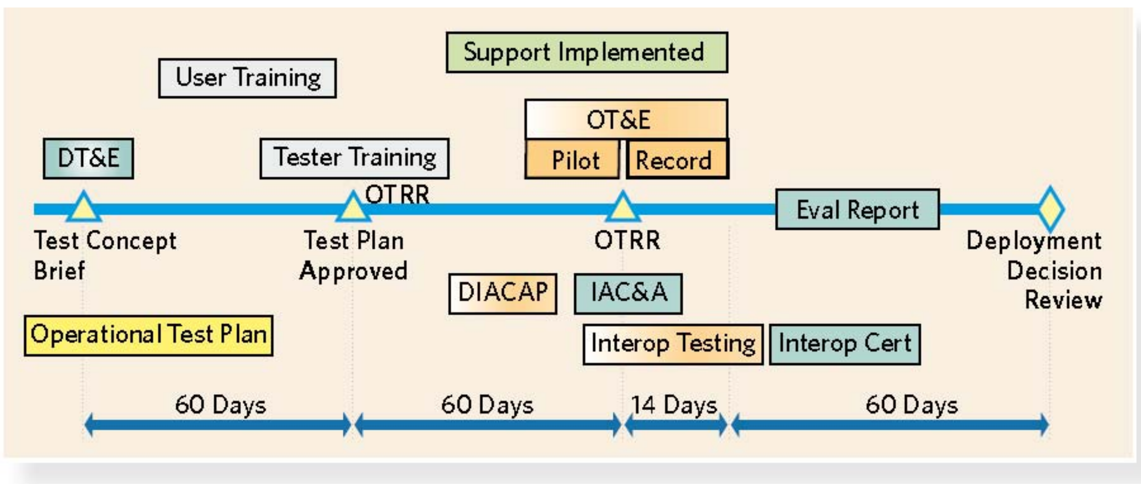

Figure 2. Test Execution Window
(Hutchison, 2010, p. 25)

Hutchison (2010, p. 24) wrote that his model allows considerably more than enough time to reach a deployment decision review versus the actual time a program takes in the testing communities. Testing factors such as acquisition leaders, cost, and resources play a major part in how long or short a program window can be within a T&E acquisition window.

This chapter provided background for understanding the acquisition process as well as the aspect of DoD testing. The next chapter will introduce the BTA’s testing process on business systems as well as introduce data that will be analyzed in Chapter IV.

III. BUSINESS SYSTEMS

A. OVERVIEW

This chapter provides information on the BTA business system, in particular the Standard Procurement System, which is critical to the warfighters' abilities to procure goods and services. Also, it presents data that will be later analyzed in Chapter IV to address the issues that were presented in Chapter I.

The data represented in this section are provided from multiple sources: the System Procurement System (SPS) System Acceptance Tests, SPS Bug System Reports (SPS-B), SPS Customer Service Data, Congressional Research Service reports, GAO reports, and (BCL) data from the BTA.

B. BUSINESS TRANSFORMATION AGENCY

The DoD sought to develop an agency to lead business transformation efforts throughout the DoD to improve the warfighter's capabilities at a faster pace and to provide financial accountability to the taxpayers. The BTA was created in October 2005 by the approval of the Defense Business System Management Committee (DBSMC). Although there are multiple directorates that makeup the BTA, the focus of this research is the Warfighter Requirements Directorate.

The Warfighter Requirements Directorate (WR), originally known as the Warfighter Support Office (WSO), "addresses immediate business process and business system challenges that adversely impact current operations and connects the Department of Defense's (DoD) business mission to the warfighter, identifying and addressing frontline opportunities" (www.bta.mil). As stated by the BTA, the WR staff is structured to be lean, focused, and agile, with a tight focus on a small set of critical issues. This chapter analyzes one program office within the WR: the Standard Procurement System.

C. BUSINESS CAPABILITY LIFECYCLE MODEL

The Business Capability Lifecycle Model (BCL) was developed in 2007 under the direction of the former Deputy Secretary of Defense, Gordon England. The BCL provides a new acquisition management model for business systems with the goal of getting new capabilities out to the warfighters faster. This model was years ahead of the model that the DSB proposed in March 2009 that addresses information systems. The BTA noted that Directive-type Memorandum (DTM) 08-020, "Investment Review Board (IRB) Roles and Responsibilities," and Chairman of the Joint Chiefs of Staff Instructions (CJCSIs) have been instrumental in implementing the BCL.

1. DoD Investment Management Governance Evolution

As depicted by the BTA BCL overview slide shown in Figure 3, DoD business systems fell under four different processes, all of which included purpose, governance structure, and documentation requirements. These four processes were the Joint Capabilities Integration and Development Systems (JCIDS); Planning Programming, Budget, and Execution (PPBE); the Defense Acquisition System (DAS); and the Investment Review Board (IRB)/Defense Business System Management Committee (DBSMC). The DoD recognized in 2007 that in order for the BTA to perform its mission of pushing capabilities out to the warfighter faster and to be held accountable, it would need to merge all the processes except for PPBE, as depicted in Table 3, which is located in the Summary section of this chapter. Figure 4 describes the BCL Governance as a governance model that applies the principles of tiered accountability by assigning responsibilities and decision-making to the lowest level.

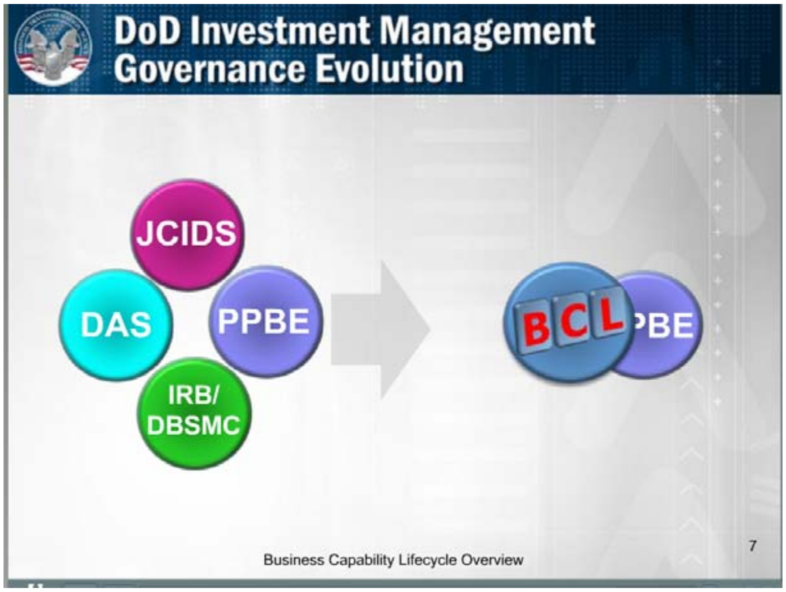

Figure 3. DoD Investment Management Governance Evolution
(Business Transformation Agency [BTA], 2009, Slide 7)

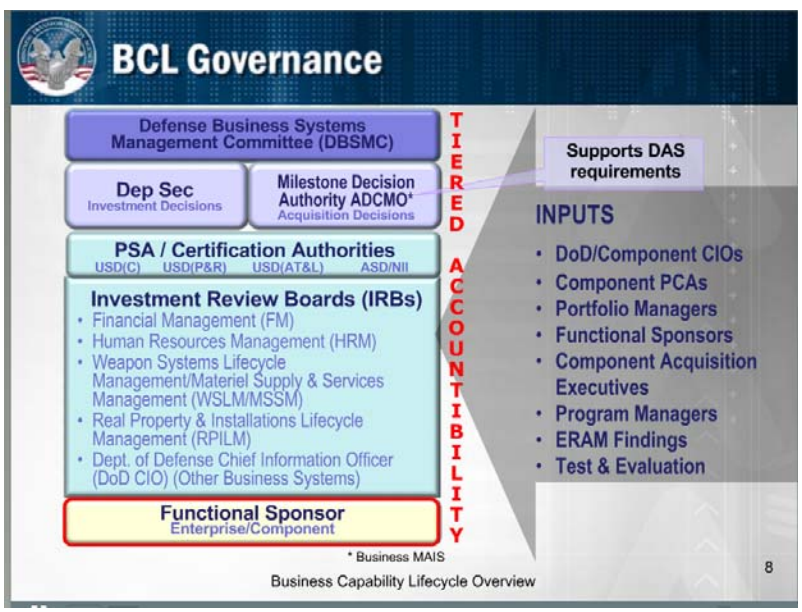

Figure 4. BCL Governance
(BTA, 2009, Slide 8)

2. BCL Management Model

The BCL management model is an evolutionary approach to acquisition. This model is made up of three phases: Business Capability Definition, Investment Management, and Execution. The BCL is an increment-based time line that facilitates program development and implementation.

The BCL functions are identified in Figure 5, which depicts the functions that occur in each phase of the model. The first phase begins with conducting an analysis and ends at Milestone A. The second phase starts with conducting solution analysis before moving to Milestone B. The final phase begins with executing the program with a Milestone C decision on whether to continue or end the program. The Enterprise Risk Assessment Methodology (ERAM) assessment is another tool that is utilized within the BCL; it is an independent assessment that can be conducted at any point within the lifecycle model. ERAM assessments are generally conducted prior to any MAIS Milestone A and B decisions. ERAM was noted as being no different from the DoD 5000 process. One of the goals of ERAM is to provide leaders with the ability to respond to emerging technology, make better decisions about how to manage program investments, and deliver business capabilities faster (Ketrick, 2009, p. 46).

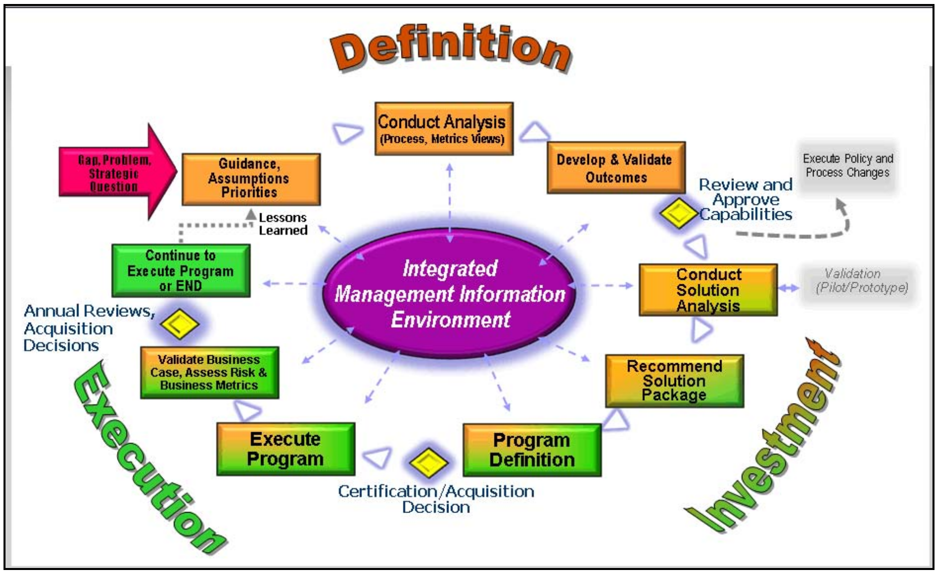

Figure 5. BTA BCL Model
(BTA, http://www.bta.mil/products/bcl.html)

D. STANDARD PROCUREMENT SYSTEM BACKGROUND

The Standard Procurement System (SPS) was established in 1994 to eliminate numerous legacy systems that conducted contract management functions. The system, over its lifecycle, has experienced some dark days with the Government Accountability Office (GAO). GAO-02-392T stated,

> this "all or nothing" approach to investing in large system acquisitions, like SPS, has repeatedly proven to be ineffective across the federal government, resulting in huge sums being invested in systems that do not provide commensurate benefits. (Government Accountability Office [GAO], 2002, p. 7)

Furthermore, the system has received recognition as it matured within its lifecycle. The Business Transformation Agency (2009) stated that the SPS program provides modern automation tools to the contracting community, which allows the procurement community to provide product and service to the warfighter on time and at reasonable prices. Currently, SPS is in the sustainment phase of its lifecycle.

E. SOFTWARE TESTING

Software testing has been a challenge for decades. Understanding what to test and how much to test plays a role in the program and has implications on cost, schedule, and performance. These are the fundamental issues that program or product managers face every day. These items are measured in multiple ways; one such way is earned value management. Earned valued management is a technique for measuring the program status as

- behind scheduled and over budget,
- ahead of schedule and over budget,
- behind schedule and under budget, or
- ahead of schedule and under budget.

It is safe to say that all program and product managers would like to be in the final category. However, this is not the case for the most part. Programs typically fail

because they do not get the requirements right the first time (along with other variables). Hence, if the requirements are not correct, it could lead to a domino effect on the cost, schedule, and performance. Cerpa and Verner (2009) cited some interesting data from the 2007 CHAOS report:

> In 2007, the Standish Group reported that 35% of software projects started in 2006 were successful compared with only 16% in the corresponding 1994 report; however, the 2007 CHAOS report still identifies 46% (53% in 1994) of software projects as “challenged” (having cost or time overruns or not fully meeting user requirements) and 19% (31% in 1994) as outright failures. (p. 130)

GAO-10-1059T stated that “according to DOD officials, the department relies on about 2,080 business systems, including accounting, acquisition, logistics, and personnel systems, to support its business functions” (GAO, 2010, p. 1). The report further goes on to mention prior work conduct by the GAO and Army Test and Evaluation found that delays in implementing certain Enterprise Resource Planning (ERPs) efforts within the DoD were due to inadequate requirements management, system testing, and data quality issues (GAO, 2010, pp. 18–19).

There are multiple types of testing that are done for DoD business systems, such as black box testing, white box testing, unit testing, integration testing, incremental integration testing, functional testing, end-to-end testing, acceptance testing, and load testing, to name a few. Defining how much to test and when to stop testing in order to move forward typically depends on the resources that are available for that particular testing activity. The data for analysis in Chapter III comes from the BTA program offices and the JITC system acceptance test report.

F. SPS DATA METHODOLOGY

In analyzing the SPS test report and SPS-B data, the researcher used data analysis techniques from *Systematic Software Testing* by Rick D. Craig and Stefan P. Jaskiel (2002). Although there are multiple models or risk matrices that can address defect issues, there is no one-fits-all model that can eliminate all bugs within a reasonable time to ship. Boris Beizer, who is a software engineer and author and who is well known in

DoD, was quoted in *Systematic Software Testing* as saying, “There is no single, valid, rational criterion for stopping. Furthermore, given any set of applicable criteria, how each is weighted depends very much upon the product, the environment, the culture and the attitude to risk” (as cited in Craig & Jaskiel, 2002, p. 264).

The analysis in this paper covers the measure of the test effectiveness in releasing a product for shipment. The data analysis will be derived from examining data from System Acceptance Test (SAT) and by also looking at Build 4, Build 5 Product, Build 5 Government Validation and Verifications (GV&V), Build 5 Regression, and SATRC02 data from the perspective of areas with the most product issues. The formulas used in this research are as follows:

- Defect Removal Efficiency = Number of Bugs Found in Testing/Number of Bugs Found in Testing + Number Not Found; and
- Defect Spoilage = (Sum of Number of Defects * Discovered Phase Age)/Total Number of Defects.

The data analysis in this research also covers the usage of what-if scenario functions embedded within the Microsoft Excel program. This portion of the analysis reviews the data from different events occurring during particular phases within the given data.

G. CUSTOMER SERVICE DATA

Customer service data play an integral part in program success. Help desk services are established in order to ensure that the customers are able to perform their daily missions. The two functions that are clearly seen from a help desk perspective are resolving customer issues and collecting data on system issues. This analysis will not address SPS customer service from the perspective of whether it is a user issue or a software issue that was not addressed or caught during testing. The customer support data lack the historical data required to conduct a thorough analysis in this paper. Therefore, the data that is reviewed comes from Table 3, which displays a sample of the data that was used in this analysis.

H. SUMMARY

This chapter provided some background information on the BTA and SPS program as well as described the data and techniques used in Chapter IV for data analysis. There are copious amounts of data that date back to September 2005–2007, consisting of Increment 3 Build 5, Build 5 GV&V, Build 5 Regression Test, and SAT.

Table 3. Bugzilla Data
(Program Manager, Standard Procurement System [SPS], 2010)

bug_id	creation_ts	Version	component	final_score	Resolution
2226	16-Sep-05	Inc 3 Build 5	Deployment	3	DOCUMENTATION
2227	17-Sep-05	Inc 3 Build 5	PD2		CLOSED
2228	19-Sep-05	Inc 3 Build 5	PD2		CLOSED
2229	19-Sep-05	Inc 3 Build 5	Deployment	2	PRODUCT_DEFECT
2242	24-Oct-05	Inc 3 Build 4 Post	PD2		DEFERRED
2243	24-Oct-05	Inc 3 Build 4 Post	PD2		DEFERRED
2244	08-Nov-05	Inc 3 Build 4 Post	PD2		DEFERRED
2245	28-Nov-05	Inc 3 Build 4 Post	PD2		DEFERRED
2263	28-Feb-06	Inc 3 Build 5	PD2	2	PRODUCT_DEFECT
2264	28-Feb-06	Inc 3 Build 5	PD2	2	PRODUCT_DEFECT
2265	28-Feb-06	Inc 3 Build 5	PD2	1	DOCUMENTATION
2266	28-Feb-06	Inc 3 Build 5	PD2		UNFUNDED_REQUIREMENT
2267	28-Feb-06	Inc 3 Build 5	PD2		CLOSED
2268	28-Feb-06	Inc 3 Build 5	PD2		CLOSED
2269	28-Feb-06	Inc 3 Build 5	PD2	3	PRODUCT_DEFECT
2270	28-Feb-06	Inc 3 Build 5	PD2		CLOSED

THIS PAGE INTENTIONALLY LEFT BLANK

IV. DATA ANALYSIS

A. OVERVIEW

With the use of Microsoft Excel, this chapter analyzes the 2,805 issues reported in the SPS-B database presented in the previous chapter. Prior to discussing the Bugzilla data that are analyzed in this chapter, it is first necessary to explain the failure definitions and scoring criteria that are assigned to each issue placed in the SPS-B and the Data Authentication Group (DAG) process. Risk-based testing is introduced briefly in order to explain how effective it can be in testing software. Defect formulas are used to determine how many defects customers are potentially receiving after the release of the last software test. In the latter portion of this chapter, the researcher reviews the proposed IT models to see how they might decrease the time it takes to get technology to warfighters.

B. BUGZILLA DAG PROCESS

Prior to any issue being placed within the database (as stated in the SPS V4.2 Increment 3 Build 5 System Acceptance Test Report; Program Manager, Standard Procurement System, 2007, it had to follow this process:

Step 1: Issue discovered.

Step 2: Issue replicated.

Step 3: Issue put into SPS-Bugzilla.

Step 4: Issue verified by Test Site Manager (TSM) and made “DAG Ready.”

Step 5: Issue sent to DAG.

Step 6: Issue scored at the Scoring Conference.

1. DAG Members

The DAG members consist of a designated Service representative, a vendor representative, and a Joint Program Management Office (JPMO) subject-matter expert SME. The JPMO SME was tasked with validating and categorizing the issues. The data presented in this paper depict the DAG categories for scoring, which categories are functional, technical, and integrations.

2. Issues Definitions

The issues scored during the DAG process were as follows: failed requirements, product defects, documentation, and training. All other issues were not scored and were listed as follows: unfunded requirements, intermittent, duplicate, and closed. The following list was taken from the SPS System Acceptance Test Report and describes how the issues were categorized (Program Manager, Standard Procurement System, 2007, p. 41).

Failed Requirements
An issue labeled *failed requirement* meant that the issue was determined to have failed a contractually cited requirement or failed something the PCO determined to be a requirement.

Product Defects
An issue labeled *product defect* meant that the software was not working properly, but it did not have a contractual requirement that specifically failed.

Intermittent Issues
If an issue was found more than once but could never be reliably and consistently replicated, it was labeled *intermittent*.

Training Issues
If an issue could be remedied by providing a solution in training materials, the disposition was labeled *training*.

Documentation Issues
If the issue was determined to be fixable by a script or operational scenario edit, the disposition was labeled *documentation*. If the vendor-delivered instructions (e.g., the installation guide or online help) required editing, the issue was similarly dispositioned *documentation*. The issue was not closed.

Unfunded Requirements
If the software currently had no contractual requirement to perform an action that possibly called for a new requirement to be written, the issue was dispositioned *unfunded requirement*.

Closed Issues
Issues that were working as designed, that could not be replicated, or that were tester errors were labeled *closed*. In some cases, duplicate issues (e.g., issues already reported and tracked through the DAG process) were identified as *closed*.

Duplicate Issues
When an issue was agreed upon by all members of the DAG to be a duplicate of another issue being tracked, it was dispositioned *duplicate issue*.

3. Failure Definition and Scoring Criteria

The failure definition and scoring criteria for the SPS-B data that is analyzed in this chapter originated from the SPS Test report. The scores are listed from one through five, with one being *mission critical* and five being *not critical*. Table 4 describes the priority the issue was assigned if it met a certain criteria. The lower the numeric value assigned to the issue, the higher the risk associated with the issue.

Table 4. Failure Definition and Scoring Criteria (Program Manager, Standard Procurement System, 2007)

	Applies if problem could …
1	a. Prevent the accomplishment of an essential capability. b. Jeopardize safety, security, or other requirement designated *critical*. Example: Cannot create or process core Procurement documents.
2	a. Adversely affect the accomplishment of an essential capability, and no work-around solution is known. b. Adversely affect technical, cost, or schedule risks to the project or in lifecycle support of the system, and no work-around solution is known. Example: Cannot create attachments or supporting documents.
3	a. Adversely affect the accomplishment of an essential capability, but a work-around solution is known. b. Adversely affect technical, cost, or schedule risks to the project or to lifecycle support of the system, but a work-around solution is known. Example: Data expected to pre-fill using copy functions does not auto populate, but user can manually enter data.
4	a. Result in user/operator inconvenience or annoyance, but does not affect a required operational or mission-essential capability. b. Result in inconvenience or annoyance for development or personnel, but does not prevent the accomplishment of the responsibilities of those personnel Example: Cannot populate a field via Favorites functionality but can use alternate-designed methods to populate.
5	Any other effect. Example: Misspellings that are not critical to the understanding of data or data transfer.

C. RISK-BASED TESTING

There are numerous software models that depict how to determine the amount of testing required before issuing software to users. Testing for bugs is an exhaustive process. Bugs will always be within any product. However, identifying and fixing these bugs by prioritizing them can be beneficial to the program office. Interaction between the designer, the tester, and the user is invaluable throughout this process, which will determine the outcome of the software. Figuring out how much to test and when to stop testing can be determined by creating a baseline for similar projects to follow. However, every model is only as good as the information embedded within it.

1. Defect Prediction Model

The defect prediction model is a simple tool that enables the developers, engineering teams, and testers to discover how many defects will be released to the user. Figure 6 can help in determining where the focus should be throughout the testing phases. This model describes the flow of defects throughout the testing phases, along with the defects the users will inherit.

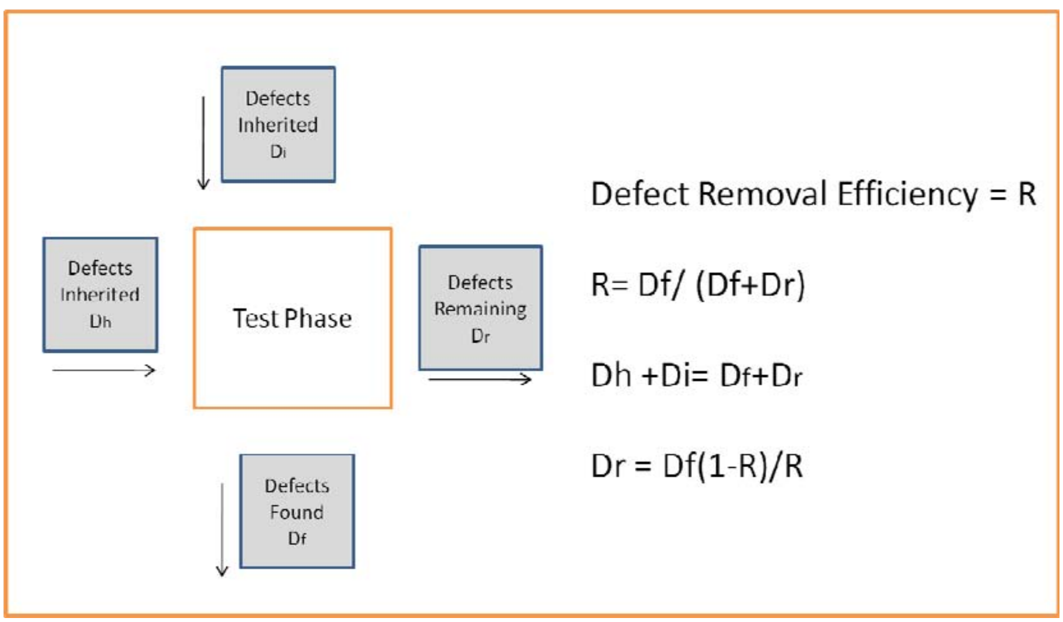

Figure 6. Defect Prediction Model

The SPS-B data is depicted within Figure 7, demonstrating the volume of defects moving from test phases to the customer. This figure also depicts a 28% defect-removal efficiency rate after the last test is conducted. This data portrays all the SPS-B data that are not labeled *closed*.

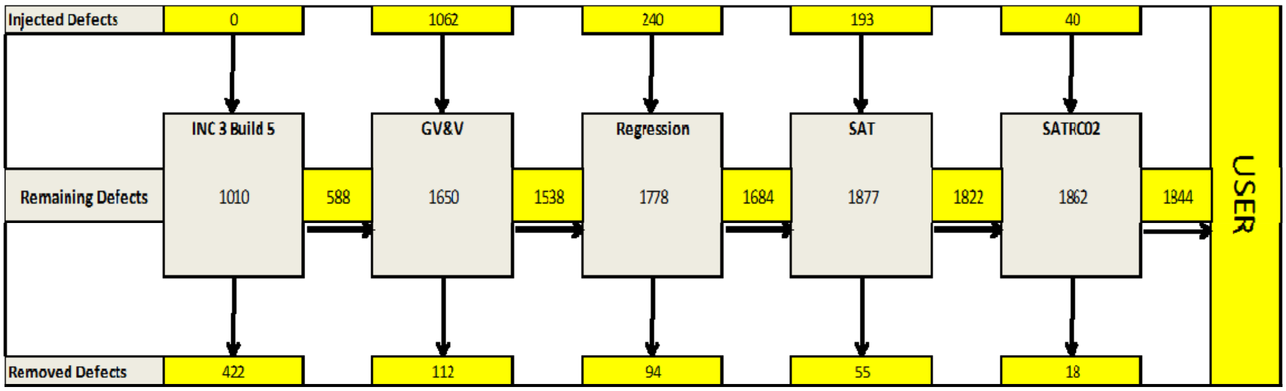

Figure 7. Increment 3 Build 5 Prediction Excel Model

The Standard Procurement System went through five testing events during which the SPS-B collected the issues. The data was populated within the SPS-B database covering the components that were tested. The components that were analyzed with the defect prediction model are the PD2, Technical, and Legacy components of the SPS. This time, the data was run from the perspective of defects that the customers will receive after the last test. The defects fell into the categories of training, documentation, product defects, and failed requirements. Table 5 depicts the total defects found in each component during the different testing phases, and Table 6 depicts the number of defects assigned a numeric value/critical value during each phase.

Table 5. Total Defects in Components

	PD2 Total Defects	Technical Total Defects	Legacy Total Defects
Inc 3 Build 5	299	Not tested	Not Tested
GVV	310	52	66
Regression	51	6	38
SAT	33	51	19
SATRC02	4	12	0

Note. The numbers in this graph were calculated from the SPS-B Excel file.

Table 6. Final DAG Scoring Numbers

	Final Scores				
	1	**2**	**3**	**4**	**5**
Inc 3 Build 5	23	196	59	19	2
GVV	112	346	136	34	0
Regression	9	45	30	11	0
SAT	13	48	37	5	0
SATRC02	0	7	6	3	0

Note. The numbers in this graph were calculated from the SPS-B Excel file.

The SPS-B data indicate, for the most part, a steady decline in defects from the beginning of the test event to the last test event, as depicted in Figure 5. Figure 6 also shows a decreasing trend on the critical aspect of the defect as it moves from one phase of testing to the next. The data that is not depicted is the data indicating where each component falls in the categories of failed requirement, product defect, documentation and training. A majority of the critical defects fell under PD2 and Legacy. The PD2 and Legacy defects were mostly found under failed requirements and product defect. Another view, from a tester perspective, is how the number of defects can be reduced from the vendor testing to GV&V. By understanding the testing requirement and user

requirement, the elapsed time for discovering additional defects during GV&V can be reduce by actively engaging the Vendor testing procedures.

D. SOFTWARE TEST COST

Software testing is extremely costly to programs and can affect the program likelihood of success in acquisition. Software re-works due to lack of requirements from users or misunderstandings from an engineer can be devastating to a program that is already over cost. The formula used in this research analysis considers the overall defect cost. Software defect cost can be described simply by using the following formula:

Software Defect Cost = Industry Salary Standards * (Number of Defects * Number of Hours Required to Fix the Defect).

This formula provides an estimate on the additional cost the program will incur with defects. Typically, if defects are caught early, cost will be low. However, when defects are caught after the release of the software, cost will be notably higher than fixing the problems early in the testing phases. This analysis used the estimated salary of software engineers, developers, and programmers with 9 years experience and a $43.99 hourly pay from http://www.payscale.com. This analysis also assumed three software engineers working 10-hour periods to fix one bug/defect. An analysis of the 1,844 defects from Figure 7 that are passed on to the customer for a quick release of the product will cost $811,175.60 to fix. This number represents labor only and would vary considerably depending on the type of contract used.

E. LIFECYCLE MODEL FOR IT SYSTEMS

The Business Capability Lifecycle Model (BCL) represents an approach to pushing equipment out to the user within 12–18 months, but faster than 24 months. The DSB–IT model shown in Figure 8 depicts an approach to reduce the time of pushing information systems out to the user with similar time frames as the BCL model. Figure 9 shows that the model Hutchinson discussed is where a certain capability is built, and then that capability is tested and sent out to the user community. This same model depicts an incremental approach after the release of the first increment. All of these models have the

same intended goal of 18 months of procuring, testing, and deploying the product to the warfighter. Systems that are tested in smaller capability increments and that are fielded earlier with planned follow-on capability upgrades provide more benefits than systems that are fielded after the technology has become antiquated. This method of building a little, testing a little, and sending the technology out to the warfighter is becoming a reality in staying current and relevant in the world. The DSB-IT and the National Research Council (NRC) have recognized this for years. “In 2006, the National Research Council observed that DoD is fast approaching a period in which a single all encompassing large-scale operational test, as currently practiced, will cease to be feasible” (Hutchison, 2010, p. 23).

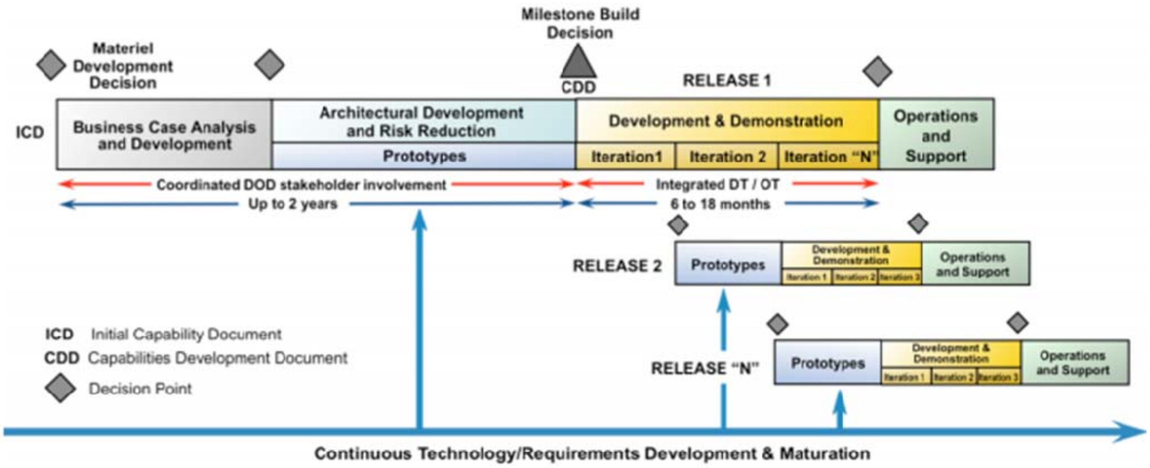

Figure 8. DSB–IT Proposed IT Acquisition Process
(DSB–IT, 2009, p. xi)

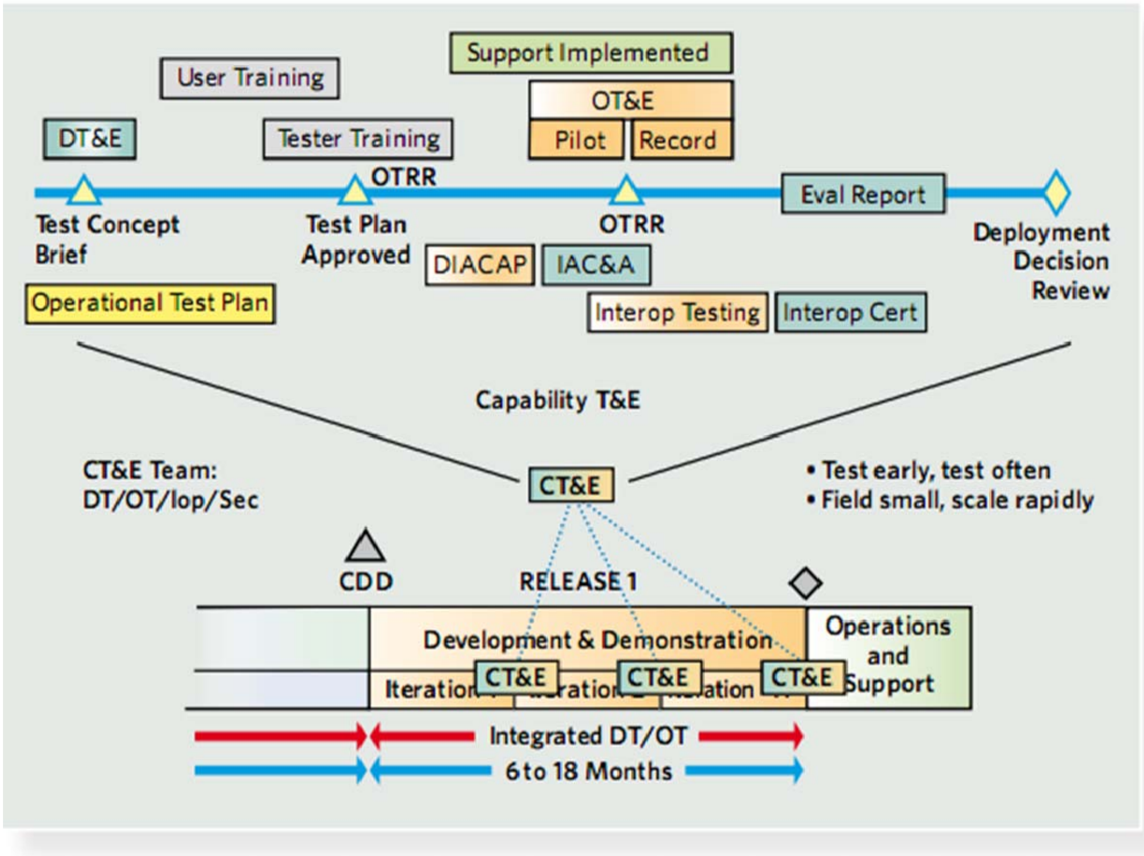

Figure 9. Agile T&E Model
(Hutchinson, 2010, p. 27)

The NRC provided a 2009 study to the DISA on how to achieve effective acquisition of technology for the DoD. The NRC stated, "The discrepancy between DoD fielding cycles and COTS life cycles is stark, and measured in years" (National Academies Press, 2010, p. 5). They went on to add the following:

- The oversight process focuses too much on shortcomings of COTS products and services and inhibits the timely delivery of meaningful (albeit imperfect) end-user capabilities.
- IT program requirements are often written with overly detailed specifications, take a long time to develop, and are not consistent with the pace of technological change or the rapid delivery of end-user capabilities.

- DoD acquisition, budgeting, requirements processes, which are designed for large weapons systems acquisition programs, are being inappropriately applied to relatively low-dollar IT programs.
- Dollar thresholds are used to assign the level of oversight for IT programs. These levels are significantly lower than the dollar levels used for determining oversight levels for weapon system programs. This disparity subjects too many IT programs to time-consuming, high-level DoD oversight and prevents the delegation of oversight to lower levels that are more agile.
- The DoD's acquisition training curriculum does not adequately address the special challenges of IT system acquisition or prepare program managers to run IT programs effectively. This shortfall impedes the DoD's ability to assess, adapt, and adopt applicable commercial methods, processes, products and services. (National Academies Press, 2010, p. 5)

Information technology can be unique and very complex; however, those risks, as they pertain to cost, schedule, and performance, can be reduce by using multiple matrices. Matrices that are non-complex in nature and that are easy to update can benefit an organization if used correctly. A defect matrix can be one of those multiple matrices that narrow the focus on were the risk is located and on how much additional effort is required to appropriately address that risk.

Each of these models represents processes that create time delays due to regulatory requirements before passing through one acquisition phase to another. In order to have an agile process that pushes capabilities out faster, the DoD requires methods that reduce the number of requirements while ensuring systems are meeting the following criteria: suitable, effective, interoperable, and secure for the warfighters. In the article by Cloutier and Crowe (2009) about fielding capabilities while using an agile process, they stated, "Our Agile and flexible approach to systems and software engineering allowed us to capture the true essence of rapid prototyping and capability deployment while still meeting budgetary, schedule, and customer satisfaction goals" (p. 17).

V. RECOMMENDATIONS AND CONCLUSION

A. PRIMARY RESEARCH QUESTION

On an ACAT III IT program, how much testing is required to ensure that risk is minimized to an acceptable level?

In determining how much testing is enough for the ACAT III IT system, the author had to determine the underlying causes that bring this question up for discussion. As noted throughout this research, the underlying cause of information technology being delayed to the warfighter is the way in which the DoD is currently controlling the development of IT systems. Although studies have concluded that there is a need to establish new information technology processes, systems will continue to be delayed to users under the existing DAS model. The rationale behind this is simple: changing the weapons system acquisition process will take more than creating a new process to follow. It will take leaders implementing these new changes with a JCIDS process that is more balanced in order to support information technology systems that are not large weapon systems. GAO-08-1060 reported that the JCIDS process had proven to be lengthy, "taking on average up to 10 months to validate a need—which further undermines efforts to effectively respond to the needs of the warfighter, especially those that are near-term" (GAO, 2008b).

The risk-based testing approach, which was discussed in Chapter IV, can minimize risk throughout the entire testing process if done correctly. In order to reduce the testing time, organizations must understand the user requirements as well as the testing requirements of each organization. It is important that during each testing phase the test community identifies critical issues and focus areas for the next phase of testing. The key to decreasing time in testing is linking the test designs back to the customer requirements. Understanding the testing scope within each organization and how integrating testing can improve the overall testing of the system will benefit all the stakeholders involved. The program manager, users, and testers working as an integrated team throughout the testing process will enable them to identify and link issues/defects

back to the critical areas that must be addressed for a successful system. This way, testing is reduced and minimized at an acceptable level. Therefore, on an ACAT III IT system, the amount of testing that is required to ensure that risk is minimized to an acceptable level is dependent on how effectively the system testing is managed.

There are numerous reasons why software systems fail to be delivered to users on time. Those reasons come in the form of the IT delivery date impacting the developer's process, actual project cost and time being underestimated, risks not being re-assessed or controlled throughout the process, and changes to software configurations being inadequately controlled throughout the lifecycle. Understanding the critical role of defect management and adequately reporting those risk and changes, along with validating those defects that are actually resolved will reduce the testing time line.

Improvement of the software process during testing can greatly help the program office in achieving a quality product prior to user acceptance. Simple, noncomplex matrices that identify bug issues and where they are within the product can be more beneficial then explaining a complex chart. Bug-capturing matrices and the early identification of issues can greatly decrease the time and cost of entering a product into its final testing phase. The impacts of using defect-removal activities have benefits that can help similar programs to determine what to emphasize during testing. This could also help them to meet testing objectives. Testing software is not effective if those doing the testing do not know what to test or how much to test.

B. SUBSIDIARY QUESTIONS

How does previous software testing such as GV&V testing, regression testing, and system acceptance testing compare and impact the final decision to deploy a system to users?

Software testing through all phases can negatively impact the next phase if rigorous testing is not conducted. As noted in Chapter IV, SPS testers and users are invaluable in the testing process. Testers and users must be integrated with software engineers during testing to ensure the system is actually working as intended in an

operational environment. Although the prediction model in Chapter IV depicted 1,844 bugs being handed over to the users, the component defects decreased throughout the components' testing phases.

Previous software testing can establish a baseline for future system testing. By understanding where the most critical issues occurred in previous testing, testers can identify and mitigate defects in future testing. Critical defects will typically decrease as more testing is conducted; however, as depicted in Table 6 in Chapter IV, DAG critical issues tend to increase in GV&V and SAT. This increase in DAG critical issues may have occurred due to testing scenarios that were never presented in previous testing phases nor known to be a requirement from a developer/tester standpoint. Users involvement can potentially reduce defects if the user community knows exactly what they need and can articulate those requirements to the acquisition community.

How much additional testing is worth the cost and time for ACAT III programs?

One way of reducing additional testing is by incorporating integrated testing with the understanding of the output from the event. DiPetto (2009) discussed integrating testing as a way of improving the quality of the information provided to decision authorities (p. 332). He went on to say, "The challenge to T&E community is to implement robust integrated testing and change the culture to fully realize the benefits to the acquisition process and ultimately to the warfighter" (DiPetto, 2009, p. 332). Testing for business systems must still go through proper testing events (i.e., information assurance, functional, interoperability, performance, and operational acceptance testing) to ensure that the intended system can functionally operate and that it is suitable for the warfighters. However, incorporating other testing agencies and the users can reduce the testing time line if requirements are understood earlier and tested throughout the testing cycle. Although most ACAT III systems are not categorized as life-threatening systems, the degree of testing lies with the test community and the program office on the scope of the test requirements. The decision of whether to continue to test must take into account the risk of conducting more tests (meaning, is it feasible to continue testing a system that has reduced all critical defects). In order to understand how much additional testing is actually worth the additional cost and time, the test community will need to understand

what those critical key performance critera are in order to develop test scenarios that accurately test the system for capabilities that are required now. By taking this approach in building a little and testing a little, users will get exactly what they need versus waiting for a capability that is antiquated.

C. RECOMMENDATION FOR FURTHER CONSIDERATION

This researcher's recommendation for further studies includes conducting actual surveys and interviews with users, testers, evaluators, and program office personnel to determine how much testing is really required from the stand point of the community. It would also be valuable to review the testing processes of program offices and the JITC to determine if they have the resources necessary for testing and capturing defects. This author believes that capturing the testing processes of an organization can determine how effective the processes really are and how they impact the program.

LIST OF REFERENCES

Alvidrez, F., Boike, R., Brubaker, P., Buddenburg, R., Cox, B., et al. (2010, May). *Industry task force report: Industry perspective on the Future of DoD IT Acquisition*. Retrieved from http://www.ndia.org/Advocacy/LegislativeandFederalIssuesUpdate/Documents/July2010/AFEI-804-Report.pdf

Andrews, R., Coffman M., Conway M., Cooper, J., Ellsworth, B., Hunter, D., & Sestak, J. (2010, March). *House armed service committee panel on defense acquisition reform findings and recommendations*. Washington, DC: Author

Axiotis, G. (2009). A new Alpha-Omega map for acquisition test and evaluation. *International Test and Evaluation Association (ITEA) Journal*, *30*(4), 473–480.

Burton, D., Hammons, C., Lapham M. A., Schenker, A., & Williams, R. (2010, April). *Technical Note-002: Considerations for using agile in DoD acquisition*. Pittsburgh, PA: Carnegie Mellon University, Software Engineering Institute.

Business Executives for National Security (BENS). (2008, August). Business executives for national security: Defense Information Systems Agency cooperative review. Retrieved from http://www.bens.org/our-work/past-work/defense-reform/disa-review.html

Business Transformation Agency (BTA). (2009). Business capability lifecycle model overview [Presentation slides]. Retrieved from http://www.bta.mil/products/training/BCL/index.html

Cerpa, N., & Verner, J. M. (2009, December).Why did your project fail? *Communication of the ACM*, *52*(12), 130–134.

Conley, S. F. (2009). Test and evaluation strategies for network-enabled systems. *International Test and Evaluation Association (ITEA) Journal*, *30*, 111–116.

Cloutier, R., & Crowe, P. (2009, May/June). Evolutionary capabilities developed and fielded in nine months. *CrossTalk*, *4*, 15–17.

Craig, R. D., & Jaskiel, S. P. (2002). *Systematic software testing*. Norwood, MA: Artech House.

Defense Acquisition University (DAU). (2005, September). *Defense acquisition guide* (7th ed.). Fort Belvoir, VA: Author.

Defense Science Board–Information Technology (DSB–IT). (2009, March). *Creating an assured joint DoD and interagency interoperable, net-centric enterprise*. Washington, DC: Author.

Defense Science Board–Task Force (DSB). (2009, March). *Department of Defense policies and procedures for the acquisition of information technology*. Washington, DC: Author.

Department of Defense (DoD). (2009, December 17). *Defense acquisition guidebook* (Ver. 1.6). Retrieved from http://dag.dau.mil

DiPetto, C. (2009). Implementing integrated testing. *International Test and Evaluation Association (ITEA) Journal*, *30*(3), 331–332.

Garamone, J. (2010, May 26). Lynn announces task force to speed IT procurement. *American Forces Press Services*. Retrieved from http://www.defense.gov/news/newsarticle.aspx?id=59354

Government Accountability Office (GAO). (2002). *DoD's standard procurement system: Continued investment has yet to be justified* (GAO-02-932T). Washington, DC: Author.

Government Accountability Office (GAO). (2008b). *DoD's requirements determination process has not been effective in prioritizing joint capabilities* (GAO-08-1060). Washington, DC: Author.

Government Accountability Office (GAO). (2010). Financial management improvement and audit readiness efforts continue to evolve (GAO-10-10597T). Washington, DC: Author.

Herrin, R., Knodle, D., Mackenzie, D., & Stephens, R. (2008). Operational realism via net-centric test and evaluation: From concept development to full rate production and sustainment. *International Test and Evaluation Association (ITEA) Journal*, *29*(2), 147–155.

Hutchinson, S. J. (2009, January/February). A capability-focused T&E framework. *Defense AT&L*, 4–20.

Hutchinson, S. J. (2010, March/April). Test and evaluation at the speed of need. *Defense AT&L*, 21–27.

Ketrick, P. (2009, September/October). Testing the enterprise risk assessment. *Defense AT&L*, *38*(1), 46–47.

National Academies Press. (2010, March). Achieving effective acquisition of information technology in the Department of Defense. Retrieved from http://www.nap.edu/catalog/12823.html

Program Manager, Standard Procurement System. (2007, March). Test report for standard procurement system: Increment 3 build 5 integrated system acceptance test [Program test report]. Fairfax, VA, Business Transformation Agency.

Program Manager,Standard Procurement System. (2010, September). Bugzilla report from increment 3 build 5 [Program test report]. Fairfax, VA, Business Transformation Agency.

Schwartz, M. (2009, July). Defense Acquisitions: How DoD acquires weapon systems and recent efforts to reform the process. Retrieved from http://opencrs.com/document/RL34026/2009-07-10/

Schwartz, M. (2010, April). Defense acquisitions: How DoD acquires weapon systems and recent efforts to reform the process. Retrieved from http://opencrs.com/document/RL34026/

Under Secretary of Defense, Acquisition, Technology, & Logistics (USD[AT&L]). (2007, November 20). *The defense acquisition system* (DoD Instruction 5000.01). Washington, DC: Author.

Under Secretary of Defense, Acquisition, Technology, & Logistics (USD[AT&L]). (2008, December 8). *Operation of the defense acquisition system* (DoD Instruction 5000.02). Washington, DC: Author.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California